REPORT DOCUMENTATION PAGE

Form Approved OMB No. 0704-0188

1. REPORT DATE (DD-MM-YYYY)	2. REPORT TYPE	3. DATES COVERED (From - To)
20-12-2011	FINAL	01-04-2008 - 30-09-2011

4. TITLE AND SUBTITLE

Decision-Theoretic Foundations for Multi-Agent Systems

5a. CONTRACT NUMBER FA9550-08-1-0181

5b. GRANT NUMBER

5c. PROGRAM ELEMENT NUMBER

5d. PROJECT NUMBER

5e. TASK NUMBER

5f. WORK UNIT NUMBER

6. AUTHOR(S)

Zilberstein, Shlomo

7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)

University of Massachusetts
Department of Computer Science
140 Governors Drive
Amherst, MA 01003-9264

8. PERFORMING ORGANIZATION REPORT NUMBER

9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES)

Office of Grant and Contract Administration
70 Butterfield Terrace
University of Massachusetts
Amherst, MA 01003

AFOSR/NL
875 N RANDOLPH STREET
SUITE 325, RM 3112
ARLINGTON VA 22203-1768

10. SPONSOR/MONITOR'S ACRONYM(S) AFOSR

11. SPONSOR/MONITOR'S REPORT NUMBER(S) AFRL-OSR-VA-TR-2012-0657

12. DISTRIBUTION/AVAILABILITY STATEMENT

DISTRIBUTION STATEMENT A. APPROVED FOR PUBLIC RELEASE; DISTRIBUTION IS UNLIMITED.

13. SUPPLEMENTARY NOTES

The views, opinions and/or findings contained in this report are those of the author, and not necessarily shared by AFOSR.

14. ABSTRACT

This project produced efficient algorithms for planning and coordination of multi-agent systems that can cope with uncertainty and missing information. These algorithms employ new plan representations and dynamic programming techniques that can exploit heuristic knowledge and the structure of the problem to improve scalability. The project produced mechanisms that exploit randomization to improve coordination and minimize communication, and has shown how to use agent goals to develop bounded-optimal algorithms that are based on sampling. Additionally, the project produced CBDP, an efficient and scalable point-based dynamic programming algorithm for Network Distributed POMDPs, particularly suited for managing sensor network tracking tasks. A formal framework for decentralized monitoring has been developed for coordination of agents that solve components of a larger problem in a decentralized manner. These new coordination algorithms have been rigorously evaluated and shown to produce magnitudes of speedup in policy computation and better quality solutions than state-of-the-art methods.

15. SUBJECT TERMS

16. SECURITY CLASSIFICATION OF: a. REPORT	b. ABSTRACT	c. THIS PAGE	17. LIMITATION OF ABSTRACT	18. NUMBER OF PAGES	19a. NAME OF RESPONSIBLE PERSON / 19b. TELEPHONE NUMBER (Include area code)
U	U	U	SAR	8	Shlomo Zilberstein / 413-545-4189

Standard Form 298 (Rev. 8/98)
Prescribed by ANSI Std. Z39.18
Adobe Professional 7.0

Final Performance Report

Decision-Theoretic Foundations for Multi-Agent Systems

AFOSR Agreement Number FA9550-08-1-0181

Shlomo Zilberstein, Principal Investigator
Computer Science Department
140 Governors Drive
University of Massachusetts
Amherst, MA 01003-9264

December 23, 2011

Abstract

This project produced efficient algorithms for planning and coordination of multi-agent systems that can cope with uncertainty and missing information. These algorithms employ new plan representations and dynamic programming techniques that can exploit heuristic knowledge and the structure of the problem to improve scalability. The project produced mechanisms that exploit randomization to improve coordination and minimize communication, and has shown how to use agent goals to develop bounded-optimal algorithms that are based on sampling. Additionally, the project produced CBDP, an efficient and scalable point-based dynamic programming algorithm for Network Distributed POMDPs, particularly suited for managing sensor network tracking tasks. A formal framework for decentralized monitoring has been developed for coordination of agents that solve components of a larger problem in a decentralized manner. These new coordination algorithms have been rigorously evaluated and shown to produce magnitudes of speedup in policy computation and better quality solutions than state-of-the-art methods.

1 Overview

This project produced new formal models and algorithms for coordination of complex multi-agent systems, such as target tracking sensors that have different partial information about the target. Similar problems of planning under uncertainty that involve a single decision maker have been studied extensively using the partially observable Markov decision process (POMDP). We have developed an extension of POMDPs for multi-agent settings where decisions are made in a decentralized manner by a group of decision makers. The new model is called decentralized partially observable Markov decision process (DEC-POMDP). While POMDPs and DEC-POMDPs offer a rich representation for sequential decision making under uncertainty, the computational complexity of each model presents important research challenges.

To address these challenges, we have developed several solution methods based on utilizing domain structure, memory-bounded representations and sampling. These approaches address some of the major bottlenecks for decision making in real-world multi-agent systems. The methods include a more efficient optimal algorithm for DEC-POMDPs as well as scalable approximate algorithms for POMDPs and DEC-POMDPs. The optimal approach, incremental policy generation, uses reachability in the domain structure

20120918116

to more efficiently generate solutions for a given problem by focusing the search on decisions that may contribute to the optimal solution. This approach can also be used in other optimal or approximate dynamic programming algorithm for DEC-POMDPs, improving their scalability. We have also shown that approximate dynamic programming can be improved by sampling to determine areas that are likely to be visited by other agent policies. Other approximate algorithms we developed include a memory-bounded approach that represents optimal fixed-size solutions for POMDPs and DEC-POMDPs as nonlinear programs. This representation can be used to solve the coordination problem optimally as well as approximately. We have also shown that increased solution quality and scalability can be achieved by using the more compact and structured representation of Mealy machines. Mealy machines can be used to represent agent policies in any infinite-horizon approach that uses finite-state controllers, allowing memory to be utilized more efficiently and domain information to be automatically integrated into the solution representation. We have also made use of domain structure in the form of goals. We summarize these results below. Further details are available in the listed publications.

2 Summary of Research Challenges and Accomplishments

Constraint-Based Dynamic Programming for Decentralized POMDPs with Structured Interactions We have studied several ways to overcome the computational complexity of the general DEC-POMDP model by focusing on structured classes of DEC-POMDPs, which are easier to solve yet rich enough to capture many practical problems. We developed CBDP, an efficient and scalable point-based dynamic programming algorithm for one such model called ND-POMDP (Network Distributed POMDP). CBDP provides magnitudes of speedup in the policy computation and generates better quality solutions for all test instances. It has linear complexity in the number of agents and horizon length. Furthermore, the complexity per horizon for the examined class of problems is exponential only in a small parameter that depends upon the interaction among the agents, achieving significant scalability for large, loosely coupled multi-agent systems. The efficiency of CBDP lies in exploiting the structure of interactions using constraint networks.

Event-Detecting Multi-Agent MDPs Following the same principles of exploiting domain structure, we developed the class of event-detecting multi-agent MDPs (eMMDPs), designed to detect multiple mobile targets by a team of sensor agents. We have shown that eMMDPs are NP-Hard and presented a scalable factor-2 approximation algorithm for solving them using matroid theory and constraint optimization. The complexity of the algorithm is linear in the state-space and number of agents, quadratic in the horizon, and exponential only in a small parameter that depends on the interaction among the agents. Despite the worst-case approximation ratio of 2, experimental results have shown that the algorithm can produce near-optimal policies for a range of test problems.

Finite-State Controllers Based on Mealy Machines for Centralized and Decentralized POMDPs Existing controller-based approaches for centralized and decentralized POMDPs are based on automata with output known as Moore machines. We have shown that several advantages can be gained by utilizing another type of automata, the Mealy machine. Mealy machines are more powerful than Moore machines, provide a richer structure that can be exploited by solution methods, and can be easily incorporated into current controller-based approaches. To demonstrate this, we adapted some existing controller-based algorithms to use Mealy machines and obtained results on a set of benchmark domains. The Mealy-based approach always outperformed the Moore-based approach and often outperformed the state-of-the-art algorithms for both centralized and decentralized POMDPs.

Point-Based Backup for Decentralized POMDPs One effective direction for addressing the complexity of the belief space in POMDP and DECP-POMDP is the use of point-based methods, which optimize the policy for a sample of belief points. Performing point-based backup is a fundamental operation in state-of-the-art algorithms. We have shown that even a single backup step in the multi-agent setting is NP-Complete. Despite this negative worst-case result, we developed an efficient and scalable optimal algorithm as well as a principled approximation scheme. The optimal algorithm exploits recent advances in the weighted CSP literature to overcome the complexity of the backup operation. The poly-time approximation scheme provides a constant factor approximation guarantee based on the number of belief points. In experiments on standard domains, the optimal approach provided a significant speedup (up to 2 orders of magnitude) over the previous best optimal algorithm and was able to increase the number of belief points by more than a factor of 3. The approximation scheme also worked well in practice, providing near-optimal solutions to the backup problem.

Decentralized Monitoring of Distributed Anytime Algorithms Anytime algorithms allow a system to trade solution quality for computation time. In previous work, monitoring techniques have been developed to allow agents to stop the computation at the "right" time so as to optimize a given time-dependent utility function. However, these results were obtained only for the single-agent case. We have analyzed the more general problems that arise when several agents solve components of a larger problem, each using an anytime algorithm. Monitoring in this case is more challenging as each agent is uncertain about the progress made so far by the others. We developed a formal framework for decentralized monitoring, established the complexity of several interesting variants of the problem, and developed solution techniques for each one. Finally, we showed that the framework can be applied to decentralized flow and planning problems.

Achieving Goals in Decentralized POMDPs Coordination of multiple agents under uncertainty in the decentralized POMDP model is known to be NEXP-complete, even when the agents have a joint set of goals. Nevertheless, we have shown that the existence of goals can help develop effective planning algorithms. We examined an approach to model these problems as indefinite-horizon decentralized POMDPs, suitable for many practical problems that terminate after some unspecified number of steps. Our algorithm for solving these problems is optimal under some common assumptions – that terminal actions exist for each agent and rewards for non-terminal actions are negative. We also developed an infinite-horizon approximation method that allows us to relax these assumptions while maintaining goal conditions. An optimality bound was developed for this sample-based approach and experimental results showed that it is able to exploit the goal structure effectively. Compared with the state-of-the-art, this approach can solve larger problems and produce significantly better solutions.

Online Planning for Ad Hoc Autonomous Agent Teams Finally, we developed a novel online planning algorithm for ad hoc team settings – challenging situations in which an agent must collaborate with unknown teammates without prior coordination. Our approach is based on constructing and solving a series of stage games, and then using biased adaptive play to choose actions. The utility function in each stage game is estimated via Monte-Carlo tree search using the UCT algorithm. We established analytically the convergence of the algorithm and showed that it performs well in a variety of ad hoc team domains.

These outcomes are described in details in archival publications listed in the following section.

3 Publications

Note: The publications are available for download at:
`http://rbr.cs.umass.edu/shlomo/`

3.1 PhD Dissertations

1. Alan Carlin, PhD 2011. "Decision-Theoretic Meta-Reasoning in Partially-Observable and Decentralized Settings"

2. Christopher Amato, PhD 2010. "Increasing Scalability in Algorithms for Centralized and Decentralized Partially Observable Markov Decision Processes."

3.2 Journal Publications

1. F. Wu, S. Zilberstein, and X. Chen. "Online Planning for Multi-Agent Systems with Bounded Communication." *Artificial Intelligence*, 175(2):487–511, 2011.

2. C. Amato, D. S. Bernstein, and S. Zilberstein. "Optimizing Fixed-Size Stochastic Controllers for POMDPs and Decentralized POMDPs." *Autonomous Agents and Multi-Agent Systems*, 21(3):293–320, 2010.

3. D. S. Bernstein, C. Amato, E. A. Hansen, and S. Zilberstein. "Policy Iteration for Decentralized Control of Markov Decision Processes." *Journal of Artificial Intelligence Research*, 34:89–132, 2009.

4. R. Becker, A. Carlin, V. Lesser, and S. Zilberstein. "Analyzing Myopic Approaches for Multi-Agent Communication." *Computational Intelligence*, 25(1):31–50, 2009.

5. S. Seuken and S. Zilberstein. "Formal Models and Algorithms for Decentralized Decision Making under Uncertainty." *Journal of Autonomous Agents and Multi-Agent Systems*, 17(2):190–250, 2008.

3.3 Conference Publications

1. A. Kumar, S. Zilberstein, and M. Toussaint. "Scalable Multiagent Planning Using Probabilistic Inference." *Proceedings of the 22nd International Joint Conference on Artificial Intelligence* (IJCAI), pp. 2140–2146, Barcelona, Spain, July 2011.

2. F. Wu, S. Zilberstein, and X. Chen. "Online Planning for Ad Hoc Autonomous Agent Teams." *Proceedings of the 22nd International Joint Conference on Artificial Intelligence* (IJCAI), pp. 439–445, Barcelona, Spain, July 2011.

3. A. Kumar and S. Zilberstein. "Message-Passing Algorithms for Quadratic Programming Formulations of MAP Estimation." *Proceedings of the 27th Conference on Uncertainty in Artificial Intelligence* (UAI), pp. 428–435, Barcelona, Spain, July 2011.

4. A. Carlin and S. Zilberstein. "Decentralized Monitoring of Anytime Decision Making." *Proceedings of the 10th International Conference on Autonomous Agents and Multiagent Systems* (AAMAS), pp. 157–164, Taipei, Taiwan, May 2011.

5. C. Amato, B. Bonet, and S. Zilberstein. "Finite-State Controllers Based on Mealy Machines for Centralized and Decentralized POMDPs." *Proceedings of the 24th Conference on Artificial Intelligence* (AAAI), pp. 1052–1058, Atlanta, Georgia, July 2010.

6. F. Wu, S. Zilberstein, and X. Chen. "Trial-Based Dynamic Programming for Multi-Agent Planning." *Proceedings of the 24th Conference on Artificial Intelligence* (AAAI), pp. 908–914, Atlanta, Georgia, July 2010.

7. A. Kumar and S. Zilberstein. "Anytime Planning for Decentralized POMDPs using Expectation Maximization." *Proceedings of the 26th Conference on Uncertainty in Artificial Intelligence* (UAI), pp. 294–301, Catalina Island, California, July 2010.

8. F. Wu, S. Zilberstein, and X. Chen. "Rollout Sampling Policy Iteration for Decentralized POMDPs." *Proceedings of the 26th Conference on Uncertainty in Artificial Intelligence* (UAI), pp. 666–673, Catalina Island, California, July 2010.

9. F. Wu, S. Zilberstein, and X. Chen. "Point-Based Policy Generation for Decentralized POMDPs." *Proceedings of the 9th International Conference on Autonomous Agents and Multiagent Systems* (AAMAS), pp. 1307–1314, Toronto, Canada, May 2010.

10. A. Kumar and S. Zilberstein. "Point-Based Backup for Decentralized POMDPs: Complexity and New Algorithms." *Proceedings of the 9th International Conference on Autonomous Agents and Multiagent Systems* (AAMAS), pp. 1315–1322, Toronto, Canada, May 2010.

11. C. Amato, J.S. Dibangoye, and S. Zilberstein. "Incremental Policy Generation for Finite-Horizon DEC-POMDPs." *Proceedings of the 19th International Conference on Automated Planning and Scheduling* (ICAPS), pp. 2–9, Thessaloniki, Greece, September 2009.

12. F. Wu, S. Zilberstein, and X. Chen. "Multi-Agent Online Planning with Communication." *Proceedings of the 19th International Conference on Automated Planning and Scheduling* (ICAPS), pp. 321–329, Thessaloniki, Greece, September 2009. (**One of 6 papers selected for AI Journal fast tracking**)

13. A. Carlin and S. Zilberstein. "Myopic and Non-Myopic Communication Under Partial Observability." *Proceedings of Intelligent Agent Technology* (IAT), pp. 331–338, Milan, Italy, September 2009.

14. A. Kumar and S. Zilberstein. "Event-Detecting Multi-Agent MDPs: Complexity and Constant-Factor Approximation." *Proceedings of the 21st International Joint Conference on Artificial Intelligence* (IJCAI), pp. 201–207, Pasadena, California, July 2009.

15. C. Amato and S. Zilberstein. "Achieving Goals in Decentralized POMDP." *Proceedings of the 8th International Conference on Autonomous Agents and Multiagent Systems* (AAMAS), pp. 593–600, Budapest, Hungary, May 2009.

16. A. Kumar and S. Zilberstein. "Constraint-Based Dynamic Programming for Decentralized POMDPs with Structured Interactions." *Proceedings of the 8th International Conference on Autonomous Agents and Multiagent Systems* (AAMAS), pp. 561–568, Budapest, Hungary, May 2009.

3.4 Workshop Publications

1. A. Kumar, W. Yeoh, and S. Zilberstein. "On Message-Passing MAP Estimation in Graphical Models and DCOPs." *Thirteenth International Workshop on Distributed Constraint Reasoning* (DCR), Barcelona, Spain, July 2011.

2. A. Kumar and S. Zilberstein. "Message Passing Algorithms for Large Structured Decentralized POMDPs." *Proceedings of the 10th International Conference on Autonomous Agents and Multiagent Systems* (AAMAS), Taipei, Taiwan, May 2011.

3. A. Carlin and S. Zilberstein. "Decentralized Decision Making with Anytime Algorithms." *NIPS 2010 Workshop on Decision Making with Multiple Imperfect Decision Makers*, Whistler, British Columbia, Canada, December 2010.

4. F. Wu, S. Zilberstein, and X. Chen. "Point-Based Policy Generation for Decentralized POMDPs." *AAMAS 2010 Workshop on Multi-Agent Sequential Decision Making in Uncertain Domains* (MSDM 2010), Toronto, Canada, May 2010.

5. A. Carlin and S. Zilberstein. "Value of Communication In Decentralized POMDPs." *AAMAS 2009 Workshop on Multi-Agent Sequential Decision Making in Uncertain Domains* (MSDM), Budapest, Hungary 2009.

6. C. Amato and S. Zilberstein. "What's Worth Memorizing: Attribute-based Planning for DEC-POMDPs." *ICAPS Workshop on Multiagent Planning*, Sydney, Australia, September 2008.

4 Interactions and Transitions

The project team was very active in several conferences, symposia, panels, and journals. Two graduate student assigned to this project, Alan Carlin and Christopher Amato, have completed their PhD dissertations. Team members were engaged in several international collaborations. These interactions, which help disseminate the results of the project, are summarized below.

4.1 Editorial Positions

1. The PI is currently the Editor-in-Chief of the *Journal of Artificial Intelligence Research*, one of the top journals in the field of AI. He has been serving on the editorial board of the journal since 2002.

2. The PI serves on the editorial board of two other journals: *Autonomous Agents and Multi-Agent Systems* and *Annals of Mathematics and Artificial Intelligence*.

4.2 Participation in Conference and Workshop Organization

The PI and members of this project team served extensively on the program committees of the following venues.

1. Twenty-Fifth AAAI Conference on Artificial Intelligence, July 7-11, 2011, San Francisco, California.
2. AAAI 2011 Workshop on Generalized Planning, August 8, 2011, San Francisco, California.
3. Twenty-Second International Joint Conference on Artificial Intelligence, July 16-22, 2011, Barcelona, Spain.
4. IJCAI 2011 Workshop on Decision Making in Partially Observable, Uncertain Worlds: Exploring Insights from Multiple Communities, July 18, 2011, Barcelona, Spain.
5. Twenty-First International Conference on Automated Planning and Scheduling, June 11-16, 2011, Freiburg, Germany.
6. Tenth International Conference on Autonomous Agents and Multiagent Systems May 2-6, 2011, Taipei, Taiwan.
7. Twenty-Fourth AAAI Conference on Artificial Intelligence, July 11-15, 2010, Atlanta, Georgia.
8. AAAI 2010 Workshop on Metacognition for Robust Social Systems, July 11-12, 2010, Atlanta, Georgia.

9. Second International Conference on Computational Sustainability, June 28-30, 2010, Boston, Massachusetts.
10. Twentieth International Conference on Automated Planning and Scheduling, May 12-16, 2010, Toronto, Canada.
11. AAMAS 2010 Workshop on Multi-Agent Sequential Decision Making in Uncertain Domains, May, 2010, Toronto, Canada.
12. Eleventh International Symposium on Artificial Intelligence and Mathematics, January 6-8, 2010, Fort Lauderdale, Florida.
13. International Conference on Automated Planning and Scheduling, September 19-23, 2009, Thessaloniki, Greece.
14. ICAPS 2009 Workshop on Generalized Planning: Macros, Loops, Domain Control, September 20, 2009, Thessaloniki, Greece.
15. Twenty-First International Joint Conference on Artificial Intelligence, July 11-17, 2009, Pasadena, California.
16. Eighth International Conference on Autonomous Agents and Multiagent Systems, May 10-15, 2009, Budapest, Hungary.
17. AAMAS 2009 Workshop on Multi-Agent Sequential Decision Making in Uncertain Domains, May 11, 2009, Budapest, Hungary.
18. International Conference on Automated Planning and Scheduling, September 14-18, 2008, Sydney, Australia.
19. ICAPS 2008 Workshop on Multiagent Planning, September 14 or 15, 2008, Sydney.
20. Twenty-Third AAAI Conference on Artificial Intelligence, July 13-17, 2008, Chicago, Illinois.
21. The First International Symposium on Search in Artificial Intelligence and Robotics, July 13-14, 2008, Chicago, Illinois.
22. AAAI 2008 Workshop on Metareasoning: Thinking about Thinking, July 13-14, 2008, Chicago, Illinois.
23. Seventh International Conference on Autonomous Agents and Multiagent Systems, May 12-16, 2008, Estoril, Portugal.
24. AAMAS 2008 Workshop on Multi-Agent Sequential Decision Making in Uncertain Domains, May 12-13, 2008, Estoril, Portugal.

4.3 Other Interactions

1. The PI has maintained close collaboration ties between his lab and the MAIA group at INRIA, Nancy, France. To advance this collaboration, INRIA has provided funding for exchange of students and short visits. The PI has also participated in a multi-institutional NSF grant that provided additional funding for this collaboration. Additionally, the PI hosted a student, Feng Wu, from the University of Science and Technology of China who has made very significant contributions to this project.
2. The PI is currently the president of the ICAPS Executive Council, which oversees the annual International Conference on Automated Planning and Scheduling—the premier venue for researchers and practitioners in the area of planning and scheduling.

5 Inventions and Patent Disclosures

None.

6 Honors and Awards

1. The PI was elected as a Fellow of the Association for the Advancement of Artificial Intelligence (AAAI) for "significant contributions to decision-theoretic reasoning, resource-bounded reasoning, automated planning, decentralized decision making and multi-agent systems." The AAAI celebrated this honor at a Fellows dinner during AAAI-11 in San Francisco, California.